Understanding the Native-Immigrant Test Scores Gap: The Case of Finland

Marc Riudavets

Supervisor: Roope Uusitalo

Univeristy of Helsinki

May 2020

Abstract

This Master's Thesis studies the immigrant-native gap in terms of test performance in the PISA assessments for Finland. Data suggest that students with immigrant background perform worse than natives. I provide evidence that this unconditional gap is partially offset when some variables are introduced in the analysis. Socioeconomic status and age of arrival seem to play the biggest role in explaining the difference in performance. Some degree of heterogeneity is also found within the immigrant population in terms of performance across the socioeconomic status distribution or the language spoken at home.

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta – Fakultet – Faculty Faculty of Social Sciences		Koulutusohjelma – Utbildningsprogram – Degree Programme Master in Economics
Tekijä – Författare – Author Marc Riudavets		
Työn nimi – Arbetets titel – Title Understanding the Native-Immigrant Test Scores Gap: The Case of Finland		
Oppiaine/Opintosuunta – Läroämne/Studieinriktning – Subject/Study track Research Track		
Työn laji – Arbetets art – Level Master	Aika – Datum – Month and year May	Sivumäärä – Sidoantal – Number of pages 37

Tiivistelmä – Referat – Abstract

This Master's Thesis studies the immigrant-native gap in terms of test performance in the PISA assessments for Finland. Data suggest that students with immigrant background perform worse than natives. I provide evidence that this unconditional gap is partially offset when some variables are introduced in the analysis. Socioeconomic status and age of arrival seem to play the biggest role in explaining the difference in performance. Some degree of heterogeneity is also found within the immigrant population in terms of performance across the socioeconomic status distribution or the language spoken at home.

Avainsanat – Nyckelord – Keywords
Immigration, PISA, educational gap

Ohjaaja tai ohjaajat – Handledare – Supervisor or supervisors
Roope Uusitalo

Säilytyspaikka – Förvaringställe – Where deposited
E-thesis

Muita tietoja – Övriga uppgifter – Additional information

Contents

1 Introduction

According to the Organisation for Economic Co-Operation and Development (2010), Finland is one of the world's leaders in the academic performance of its secondary school students, a position that has been held for the past decade and that it is reflected in international assessments like PISA. This success is largely explain by the comprehensive school reform that began in the early 1970's. Its basic goal was to guarantee the equal opportunity to a 9-year basic education, regardless of the parents' socioeconomic status, by extending the age at which students chose the academic track they wanted to pursue. Results from this reform suggest that in Finland, intergenerational income elasticity was reduced by seven percentage points indicating, in turn, that policies that expand access to academic secondary education may significantly enhance intergenerational earnings mobility (Pekkarinen, Uusitalo, Kerr, 2009). However, as Salmela-Aro & Chmielewski (2019) state, during the recent years there has been a decline in the performance of the Finnish students in the PISA test although they have still outperformed most of the other countries tested (see Figure 1).

Regardless of this outstanding results, in Finland, similar to most European countries, there still exist a gap regarding the degree of educational attainment between those students with migratory background and natives (Schnepf, 2007; Dustmann et al., 2012). In particular for Finland, Ansala et al. (2019) found that there exist an unconditional gap between natives and immigrants students in their measures of educational attainment. However, at the same time, the authors also found that this gap is offset -specially for those students that had at least one of their parents born in Finland- once they controlled for variables such as age of migration or the neighborhood where the students live.

Despite this results, the native-immigrant gap and its consequences has still raised some con-

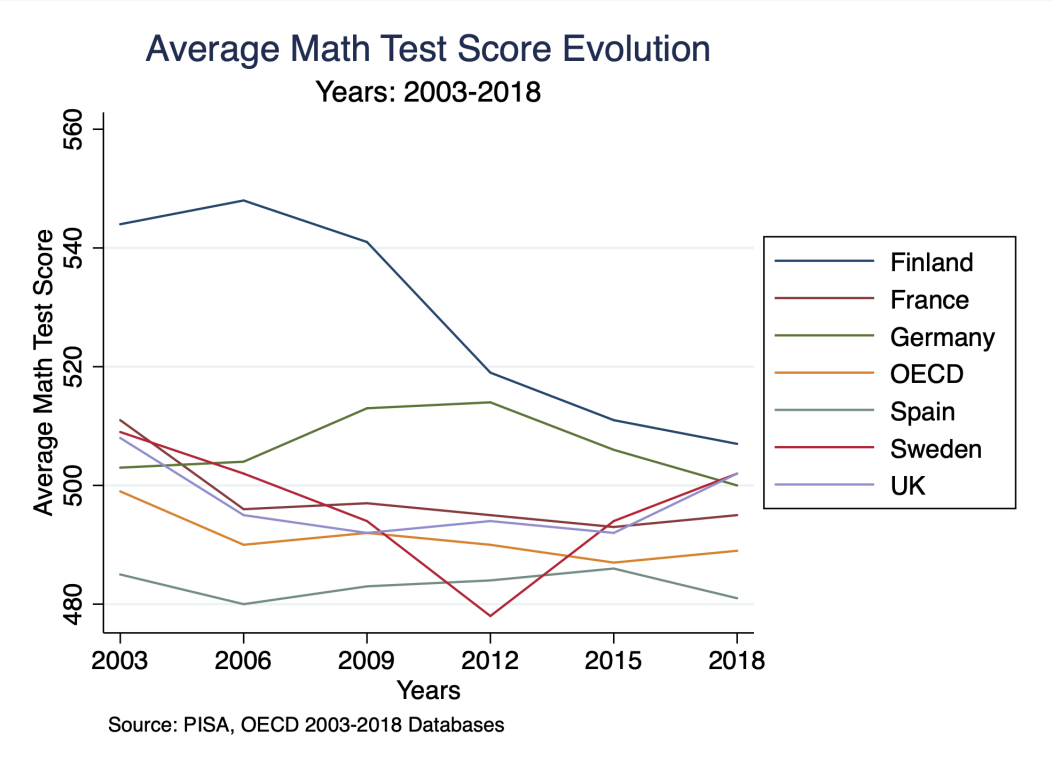

Figure 1: Average Math Test Score Evolution by Country, 2003 to 2018

cerns across Europe, becoming a central topic in most public debates and leading into a political arena where some parties are shifting towards more restrictive immigration policy proposals.

On the one hand, it is true that the existence of a native-immigrant educational gap can be an impediment to integration. Returns to schooling are an important determinant of further life time earnings and unemployment spells and hence, such divergences in educational outcomes can have a direct and lasting impact on the economic and social integration of immigrants and their descendants (Card & Schmidt 2003; Card, 2005). However, on the other hand, previous literature provides various reasons to believe that this native-immigrant gap can be partially explained beyond the native-immigrant categorisation and that the unconditional differences between natives and immigrants may be attributed to other reasons such as differences in socioeconomic background or access to educational services, among others. In fact, literature on intergenerational mobility has highlighted parental characteristics as one of the main variables to explain children´s educational attainment (Haveman and Wolfe, 1995; Dearden et al., 1997).

Focusing on the gap between native and immigrants, although Borjas (1995) pointed out that the correlation between parents skills and those of their children is not high enough to eliminate the ethnic skills differentials between native and immigrants, Cameron and Heckman (2001) found that the racial-ethnic gap in educational attainment in the US for years 1979-91 is offset or even reversed once they adjust for differences in parental socioeconomic background. Similar results are found by Van Ours and Veenman (2003) for the Netherlans, where the educational attainment of the immigrants was similar to those of the native Dutch once they took into account the education of the parents.

More recent literature on this topic has found particularly important to also include age of arrival as a key factor of the integration process. In particular, Aslund et al.(2015), by comparing siblings who arrive at different ages, or whose parents have stayed different number of years in the host country before giving birth, found that in Sweden age of arrival and hence, time spend in the country, were key factors in the assimilation process. In addition to that, Böhlmark (2018), exploiting again the within family variation for Swedish register data, found that age at immigration has a particular important negative effect on the school performance among immigrant ninth-graders. In particular, denoting the age of 9 as the critical age for which the negative performance starts to get significantly large.

The fact that age of immigration plays a significant role in determining school outcomes for young immigrants may highly the fact that there must be some knowledge that can't be transferred from the home to the host country. For this reason, another aspect which literature has been paying attention is in the role that pre-primary education and early child care attendance play in the integration process by providing the young immigrants those type of country specific skills they will further need. An aspect that could turn to be particularly relevant for Finland as compulsory school does not start until the age of 6 and, as Schneeweis (2011) showed, the

native-immigrant native gaps are smaller in host countries where school starts at an earlier age. From a general perspective, and even though literature on the topic is far from being unified, is it considered that early child care attendance positive influence further child development (Currie and Almond 2011; Ruhm and Waldfogel 2012). Moreover, if we consider that there are some type of country specific skills, such as mastering the new language, early child care attendance should be specially important due to the dynamic complementarities with future schooling stages (Cuhna and Heckman, 2007). Nevertheless, it also needs to be taken into account that the effects of pre-primary education are very heterogenous across the population and, that they have been found to be particularly important for those at the bottom of the income distirbution (Dranga & Havnes, 2019; Cornelissen et al., 2018).

This thesis then, aims to contribute in understanding how the offsprings of immigrants have performed during the recent past compared to native students. By using the PISA results and identifying the factors, if any, that this divergence in performance can be attributed, this study presents a descriptive, yet structured, discussion on the test score gap and its main drivers. One of the evident problems in this study is that immigrants are not a random sample and hence, it is necessary to account for the existence of either positive or negative selection. In addition to that, this study also tries to identify a set of observable characteristics that can be correlated with the performance of students and that are policy relevant so that it can be used in future research. In order to fulfill the aim of the study, a general picture of how the test performance gap between native and immigrants has evolved during the PISA rounds available is first presented. Then for sample size issues, the first rounds are discarded while the the 2012, 2015 and 2018 cycles are pooled together. With those, this study evaluates the unconditional gap and assess how it changes when introducing a set of different control variables that can be considered policy relevant such as the socioeconomic status, the years spend in day care or

the age of arrival. The benefit of using the PISA test scores as the main outcome variable is that it captures students at the age of 15, before completing their secondary education. The baseline analysis is complemented by examining how the gap behaves when sub-sampling first, in percentiles and secondly by language spoken at home in order to undercover the different layers of heterogeneity that may exists among the immigrant population. However, when doing so the data suffers from having to little immigrants sampled and one should take results with precaution.

Finally, this study can be a good complement to the existing literature for Finland as so far results have been focusing on outcomes such as secondary educational attainment or enrollment in higher education. Here the test scores gap is observed, which may be an indicator not only of whether they graduate or not, but up to which extend students with immigrant background perform relative to natives. Moreover, identifying correlations among different groups can be a good pointing direction on where to continue next and who should policy makers target.

The paper is structured as follows. Section 2 briefly sets in context the Finnish immigration history and its institutional background. Section 3 presents the data and the main descriptive statistics. Next, results are presented in section 4 and to finalize, section 5 sums up the main conclusions together with some possible follow-up research and policy implications.

2 Immigration in Finland

Two things should be taken into account when studding differences between native and immigrants in Finland. First, as our sample indicates, Finland immigration history is relatively short. In fact, Finland strictly restricted immigration until the early 1990s and as a result, by 1990 foreign born population was only 1.3% (Ansala et al., 2018). Its immigration history is

characterized by being a host country for Chilean refugees after the 1973 coup d'état leaded by Augusto Pinochet and, by also hosting refugees from the civil war in former Yugoslavia and Somalia. Nevertheless, during the recent years, the amount of immigrants arriving to Finland has increased, specially those coming from Russia and Estonia. Second, Finland, similarly to the other Nordic countries, presents an institutional setting that has one of the highest intergenrational mobility rates (Jänti et al., 2006). Therefore, if we take into account its short immigration history and high intergenerational mobility, policymakers in Finland have, so far, relied in universal policies, tackling low-income groups rather than specifically targeting migrants, and for the case that concerns here, children of immigrants.

Despite all the universal policies, in Finland students with immigrant background still tend to fall behind. For this reason, it is important to understand the main reasons why they don't succeed in the Finnish educational system so that specific policies, if needed, can be designed. In fact, the Finnish educational system, after its comprehensive school reform in the 1970s, is characterized by a common track and public primary and secondary education until the age of 16. However, the entry into the primary education in not mandatory until the age of 6 and some argue that this could be one of the reasons why immigrant students fall behind (Borgna, 2016).

3 Data

Data is based on the PISA databases from 2003 until 2018 for Finland. PISA is an international standardized test that assesses 15-year old students in three different domains: reading, mathematics and science. Data then, contains comparable information regarding student performance as well as a set of individual, family and background characteristics collected through self reported questionnaires. Additionally, with the data gathered, PISA constructs a set of in-

dexes that complement the already wide existing picture for each individual sampled. The main index used to control for student background characteristics in this study is the Socioeconomic Status Index (ESCS). The ESCS is constructed using three main components: measures of parental education in years, a measure of the highest parental occupation and an index, called HOMEPOS, which is constructed by asking to the students if they poses a country specific set of goods and it intends to be a proxy for family wealth as no direct measure of income is available.

PISA test is usually administrated to between 4,500 and 10,000 students in each country (Dustman et al., 2012). For Finland in particular, the amount of students who participated on the test range from 4,665 in 2006 to 8,625 in 2012. To select those students who will be assessed PISA uses a two stage sampling procedure in which first schools are sampled and then, within participating schools, students are then sampled. Afterwards, weights, provided in each data set, need to be associated to each student and school to correct for the fact that: each student or school does not necessarily have the same probability of selection, differential participation rates according to certain type of schools or student characteristics required various non-response adjustments or, some explicit strata was over sampled for national reporting purposes (OECD, 2009). For example, in 2012 immigrant population in Finland was over sampled as information about them was relatively scarce given that they represent a small fraction of the entire population. Moreover, when taking into account the two stage sampling procedure it is needed to be born in mind that students attending the same school cannot be considered as independent observations. Students within school share a set of common features that they don't do with those from a different school, like type of school, amount of resources or the quality of the teaching. Hence, a two-stage sampling procedure is less likely to cover the whole diversity of population compared to a simple random sampling procedure. Due to this, standard errors will

turn out to be larger for a two stage sample than for a random sample of the same size. For this reasons, standard errors will be clustered at school level, assuming in turn that errors will be uncorrelated across schools but that they may be correlated for students belonging to the same cluster/school.

In the analysis, the PISA definitions for immigrant status will be followed when dividing students into three categories depending on the origins of their parents. Natives are going to be those students who were born in the country where they were assessed by PISA and had at least one parent born in the country, second-generation are going to be those who were born in the country of assessment but whose parents are foreign-born and finally, first-generation are going to be those foreign-born students whose parents are also foreign-born.

The analysis starts with the entire population of students, for the 6 years I have, so that the evolution of the native-immigrant gap and its trend can be studied. As it can bee seen from Table 1, the earlier rounds, 2003, 2006 and 2009, have very few observations representing students with immigrant background. Due to this, when further pooling the sample,the analysis will be restricted to years 2012 onward. This leaves the data with a sample of 18,116 native students, 812 who are in the category of second-generation immigrants and 978 who are first generation immigrants. The main reason to restrict the sample to such years is because when aiming to compute and compare the gap between those students in the higher percentiles of the socioeconomic distribution to those in the lower percentiles of the distribution, early years just don't have enough observations to do so. Finally, when interpreting the results is it important to take into account that they will be reflecting how much difference there exists between first generation and second generation immigrants compared to natives, whom are use as baseline category.

Student performance will be measured using test scores in both mathematics and reading

Table 1: Sample Size and Outcome Variables

	Native	Second Generation	First Generation	Native	Second Generation	First Generation
		2003			2006	
Sample	5628	3	97	4601	11	61
Average Test Score Math	544.29	451.97	470.17	552.11	466.55	467.49
Average Test Score Reading	543.28	580	447.59	549.05	491.74	490.33
		2009			2012	
Sample	5624	59	71	7460	583	687
Average Test Score Math	538.85	492.95	488.07	521.72	447.29	428.18
Average Test Score Reading	532.70	495.98	458.58	526.24	458.71	417.92
		2015			2018	
Sample	5564	104	126	5175	133	179
Average Test Score Math	514.48	464.86	441.14	514.14	456.07	439.95
Average Test Score Reading	532.10	480.57	416.19	528.12	455.61	423.84

Note: This table reports the average of the outcome PISA score by immigrant categorisation.

literacy tests which can be viewed as a good predictor of future educational attainment. Test scores are reported via five plausible values for each pupil. This is because each student is tested on a randomly drawn subset of the total amount of questions and thus, five plausible values are estimated to calculate a probability distribution of the final test score for each student. Values are standardized with a mean score of 500 and a standard deviation of 100. Table 1 reports the mean test scores for both math and reading tests for all the year in my sample.

3.1 Descriptive Statistics

Table 1 shows the sample size and the average of the test scores used as outcome variables by immigrant category. As expected from the graphs of the average test score evolution in section one, a decline can be observed in the performance of the native students through the years. Moreover, there seems to be an unconditional gap in the test scores between native students and those with immigrant background, being the gap bigger for those included in the category of "first generation immigrants" from 2009 onwards. The oversampling of the population with immigrant background can be clearly observed in 2012.

Table 2 presents the variables that refer to the student background characteristics. In general we observe that the population with immigrant background have, on average, a lower socioeconomic status (SES) compared to natives. The average native student is always above the 50th percentile while those students with immigrant background, except for the second generation of immigrants on 2009, are always below the 50th percentile. Nevertheless, in Figure 2, we can also observe that, although immigrants tend to be more present in the lower parts of the SES distribution, within the immigrant populations there exists some heterogeneity. This may be an indicator that the immigrant population is not homogeneous and that gaps in test-scores may highly vary both across and within groups. This heterogeneity is also reflected when looking at the percentage distribution of the language spoken at home. A plausible feature is that those immigrants who report to speak Finnish at home are reflecting the presence of native migrants who have returned to Finland.

Table 2: Background characteristics

		Native	Second Generation	First Generation	Native	Second Generation	First Generatoin
			2003			2006	
Index of Socioeconomic Status average percentile		54.99	59	50.51	55.39	50.90	38.52
Gender (%)							
	Female	50.93	0	46.32	50.79	45.45	49.18
	Male	49.07	100	53.68	49.21	54.55	50.82
Language at home (%)							
	Finnish	82.55	33.33	15.46	93.09	18.18	19.67
	Swedish	15.87	33.33	6.19	5.76	9.09	4.92
	Other	1.58	33.33	78.35	1.15	72.73	75.41
			2009			2012	
Index of Socioeconomic Status average percentile		55.05	51.35	42.67	57.11	38.93	37.71
Gender (%)							
	Female	51.12	47.46	60	49.75	53.91	50.67
	Male	48.85	52.54	40	50.25	46.09	49.33
Language at home (%)							
	Finnish	79.94	32.20	16.90	81.38	21.78	9.61
	Swedish	19.12	3.39	2.82	15.83	1.20	1.46
	Other	0.93	64.41	80.28	2.79	77.02	88.93
			2015			2018	
Index of Socioeconomic Status average percentile		55.60	39.71	45.23	56.10	40.75	40.44
Gender (%)							
	Female	49.19	44.23	34.92	49.68	49.62	50.28
	Male	50.81	55.77	65.08	50.32	50.38	49.72
Language at home (%)							
	Finnish	94.10	24.04	18.25	92.83	33.83	16.76
	Swedish	4.70	0.96	3.17	5.53	2.26	1.68
	Other	1.2	75	78.58	1.64	63.91	81.56

Note: Average background characteristics are measured at the moment of taking the PISA test.

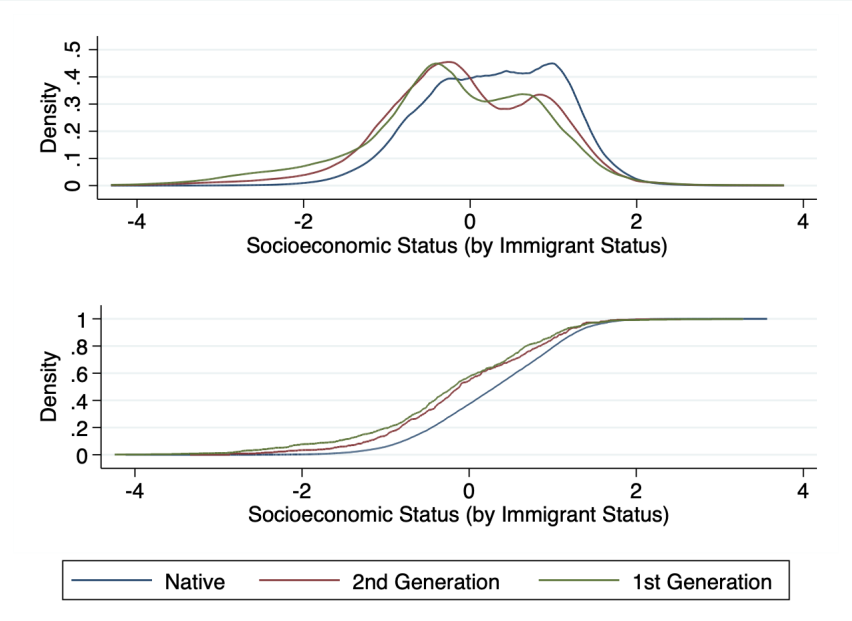

Figure 2: Socioeconomic Status sample distribution across immigrant status.
Note: Top figure plots the kernel density function and bottom one plots the cumulative density function of the Socioeconomic Status index.

Finally, Table 3 presents the percentage distribution of the variables that reflect the school characteristics. The goal of this variables is to capture those main features of the school that the students attend and, that may affect the learning process, such as the student teach ratio, the location of the school or the type of school. As expected, the Finnish educational system is mostly based on public provision which means that most of my sample, in theory, is subject to a similar education in terms of the contents taught. In addition to that, the sample of native students is more or less balanced across the location of their school, while the population with immigrant background is mostly concentrated in bigger towns and cities. Nevertheless, this concentration of immigrants in larger locations seems not generate differences in the size of the class students attend. Average size of the class is 19.4 for native students and second generation immigrants and 19.5 for first generation immigrants.

Table 3: School Characteristics

		Native	Second Generation	First Generation	Native	Second Generation	First Generation
			2003			2006	
Type of school (%)							
	Public	93.13	100	81.05	96.97	81.82	88.52
	Private	6.87	0	19.95	3.03	18.18	11.48
Location of the school (%)							
	Village	14.34	0	7.37	15.65	18.18	1.64
	Small town	32	66.67	10.53	31.22	18.18	16.39
	Town	31.71	33.33	43.16	33.57	18.18	11.48
	City	21.94	0	38.95	19.55	45.45	70.49
			2009			2012	
Type of school (%)							
	Public	95.29	88.14	92.86	94.68	87.39	94.18
	Private	4.71	11.86	7.14	5.32	12.61	5.82
Location of the school (%)							
	Village	12.94	0	12.86	6.84	0.53	0.90
	Small town	28.69	5.08	7.14	23.69	4.03	6.72
	Town	33.49	27.12	35.71	32.95	11.73	28.06
	City	24.87	67.80	44.29	36.52	83.71	64.33
			2015			2018	
Type of school (%)							
	Public	95.81	83.67	97.60	96.19	92.48	92.73
	Private	4.19	16.33	2.40	3.81	7.52	7.26
Location of the school (%)							
	Village	12.62	4.95	10.32	6.43	0.75	3.91
	Small town	24.50	2.97	11.90	25.79	7.52	13.41
	Town	36.14	29.70	36.51	36.72	24.06	16.20
	City	26.74	62.38	41.27	31.06	67.67	66.48

Note: Average school characteristics are measured at the moment of taking the PISA test.

4 Empirical Analysis

The empirical exercise conducted is based on regressing the test score outcomes to a set of explanatory variables, including as a main independent variable a dummy for the immigrant status. The immigrant dummy will capture the size of the gap. The main specification is the following:

$$y_{it} = \alpha + \beta I_{it} + \lambda X_{it} + \epsilon_{it} \tag{1}$$

where y_{it} is the test score for student i in year t, I_{it} is an indicator variable of the immigrant

status and X_{it} are variables that control for student background characteristics at individual, family and school level. Finally, ϵ_{it} is an error term clustered at school level. Time index t will be drop when pooling all the years together and, specific controls are introduced through specifications consistently with the hypothesis tested in each subsection. Nevertheless, the main parameter of interest is β, as it tries to capture the size of gap and how it evolves through the different specifications computed.

In addition to examining the test score immigrant status gap for the whole sample, the gap is also computed using a similar specification but first, across the socioeconomic status percentiles and second, for different groups based on the language they speak at home. In this cases, y is a dummy variable for the test score gap in each percentile in the first case, and a dummy variable that takes value 0 if the student speaks Finnish at home and value 1 if the students speaks the foreign language of the group analysed in the second case. In this later case, the parameter of interest, instead of capturing the immigrant-native gap, it captures the gap between those who speak Finnish at home to those who speak a specific foreign language at home.

4.1 Gap Evolution

The native-immigrant test score gap and its evolution during the recent years is first documented in Figures 3 and 4. The green line in this figures stands for the unconditional gap while the orange line accounts for the gap when the socioeconomic status is included as a control variable.

As expected, for both classifications of immigrant status a negative gap is found, meaning that immigrants perform worse than natives during the period of study. For the first generation of immigrants no pattern in the evolution of the test score gap is found (Figure 2). The larger unconditional gap is found in 2012 where the test scores for first generation immigrants is, on average, 0.94 standard deviations lower than the native students. On the other hand, the

smallest gap is in 2009 where the performance is, on average, 0.58 of a standard deviation lower. Moreover, when including the socioeconomic status as a control, aside from partially offsetting the gap, a slight increase in the importance of the socioeconomic status in explaining the test scores gap is observed.

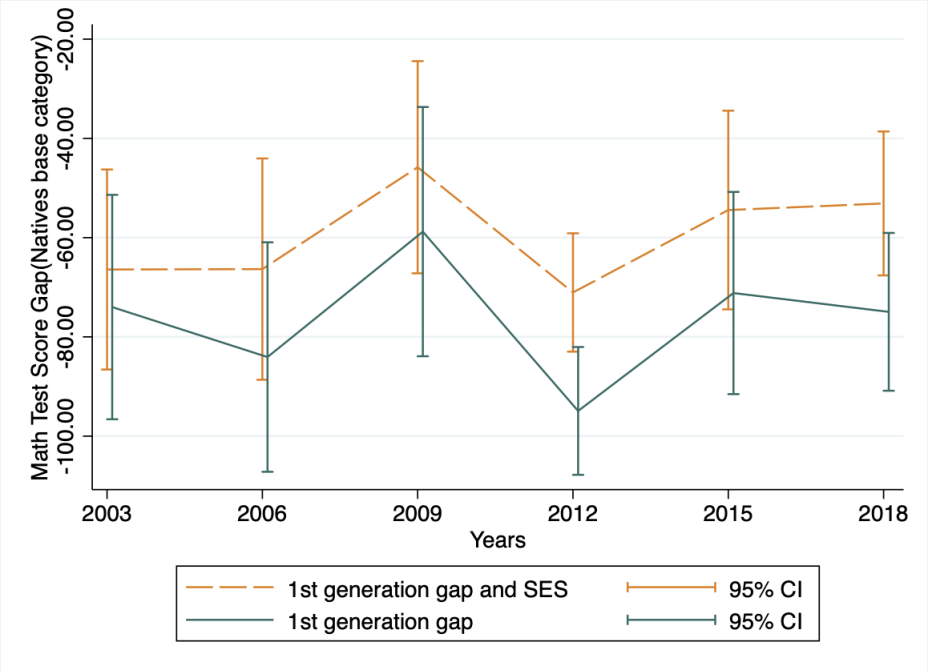

Figure 3: First generation immigrant-native gap evolution in math test scores (native as baseline category).
Note: Point estimates and 95% confidence intervals are plotted for each year. Green line presents the unconditional gap and orange line plots the gap when socioeconomic status index is included as control.

For the second generation-immigrants a smaller test score gap, compared to the one that first-generation have with the native, is observed when the first two years are not taken into account (Figure 4). The gap seems to be around 50 PISA points or half of a standard deviation. Confidence intervals for 2003 and 2006 turn out to be very large due to the fact that the sample size is very small and hence, estimates are very imprecise. Form 2009 onwards, no clear pattern can be observed in the evolution of the gap for both the unconditional and conditional specifications.

Figure 4: Second generation immigrant-native gap evolution in math test scores (native as baseline category).
Note: Point estimates and 95% confidence intervals are plotted for each year. Green line presents the unconditional gap and orange line plots the gap when socioeconomic status index is included as control.

Results from the test scores gap evolution seems to indicate that there is no closure or expansion of the gap across the years and that, despite partially explaining the gap, socioeconomic status does not account for all the difference in test performance between natives and immigrants while at same time, its importance seems to remain relatively constant across the years.

To further inspect the size of the gap and its main drivers years 2003, 2006 and 2009 are drooped and the rest of the years are pooled together. Table 4 reports the results for both the gap in the maths test and the reading literacy test. Estimates reported in columns 1 and 4 account for the unconditional gap. Columns 2 and 5 report the estimates for the specification controlling for the socioeconomic status and the gender of the student, as well as for the years fixed effects. Columns 3 and 6 add a set of school controls.

First of all it is important to note that the gap in the reading literacy test is bigger than in the math test for both first generation students and second generation students. In the unconditional specifications it is found that the gap is approximately 4 PISA points bigger for the second generation and 30 PISA points for the first generation. Assuming that the reading literacy test, compared to the math test, also accounts for a set of Finnish specific skills, results from Table 4 suggest that in general, immigrants, and specially first-generation immigrants, tend to fall behind more in subject that require country specific skills. In fact, when adding more controls, the differences in performance between the two test remain stable across the three specifications.

Table 4: Math and reading outcomes 2012-2015-2018

	(1)	(2)	(3)	(4)	(5)	(6)
	Math Test Score	Math Test Score	Math Test Score	Reading Test Score	Reading Test Score	Reading Test Score
Second-Generation	-59.65***	-41.83***	-43.48***	-63.64***	-46.07***	-50.92***
	(5.530)	(5.062)	(4.817)	(6.386)	(5.762)	(5.405)
First-Generation	-79.95***	-57.96***	-58.21***	-110.0***	-87.25***	-88.46***
	(4.885)	(4.468)	(4.521)	(5.551)	(5.186)	(5.272)
Socioeconomic Status Index		33.61***	32.35***		32.51***	30.45***
		(0.984)	(0.916)		(1.094)	(0.977)
Female		1.521	1.126		51.36***	50.74***
		(1.319)	(1.326)		(1.419)	(1.415)
Constant	517.4***	509.6***	509.1***	529.6***	492.0***	489.3***
	(1.177)	(1.929)	(7.293)	(1.325)	(2.220)	(8.567)
Year FEs	No	Yes	Yes	No	Yes	Yes
School Controls	No	No	Yes	No	No	Yes
Observations	19906	19874	19578	19906	19874	19578
R^2	0.031	0.128	0.135	0.040	0.185	0.194

Note: This table reports the math and reading test score gaps of the students who are categorised as first generation immigrants or second generation immigrants relative to native students. First generation immigrants are defined as those students who were born outside from Finland and second generation immigrants are those who were born in Finland but whose both parents were born outside of it. Models (1) and (4) reports the unconditional regressions, model (2) and (5) add controls for socioeconomic status index, gender and year fixed effects, models (3) and (6) also adds controls for school characteristics.

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Focusing now on the role that socioeconomic status plays, it can be observed that students with a 1 standard deviation higher index of socioeconomic status have approximately a 30% standard deviation higher test score gap for both the math and the reading exam. In fact, in columns 2 and 5 when including socioeconomic status as a control the gap is reduced by 30%. The further inclusion of school characteristics seems to not have any sizeable effect on the gap. If any, it looks like when including school controls the gap becomes slightly greater. A possible explanation of this is that schools in Finland are quite homogeneous across the country and hence, all students are exposed to a similar environment in terms of what school offer to them. However, in line with what Dustmant et al. (2012) argue, it is also possible that the fact that school characteristics are not important in explaining the gap might be due to measurement error in the variables included.

4.2 Years Spent in Day Care

This section examines whether differences in day care attendance can help explain part of the gap in the test scores using the cycles from 2015 and 2018 as they are the only ones that contain information regarding the amount of years students spent in day care. The motivation for examining this channel departs from the fact that Finnish compulsory education does not start until the age of 6 and understanding if pre-primary education plays a role in explaining part of the native-immigrant gap is crucial in terms of policy design issues. Clearly, attending to early child care is not a random event, for example, is it likely that there is a positive bias in terms of socioeconomic status, so no causal statements can be made.

Table 5 and 6 present the results for the math test scores gap and the reading literacy test scores gap respectively. Columns 1 first presents the unconditional gap while columns 2 includes day care attendance as a control variable. The coefficient in this specification is likely to be biased

upwards as it is probably capturing the self-selection of a certain type of students to attend day care. This intuition is confirmed in columns 3, where the inclusion of the socioeconomic status as a control variable makes the coefficient of the early child care duration reduce, although this one remains positive and significant. Finally, columns 4 includes the full set of controls for both family background and school characteristics. In this specification, although coefficients for the two different outcomes remain positive, for the math test score this one is just significant at a 5% level and it becomes not statistically different from zero in the reading specification.

Table 5: Math outcomes and and day care attendance (2015-2018)

	(1)	(2)	(3)	(4)
	Math Test Score	Math Test Score	Math Test Score	Math Test Score
Second-Generation	-54.17***	-55.65***	-34.98***	-35.19***
	(7.337)	(7.104)	(6.571)	(6.697)
First-Generation	-73.32***	-58.13***	-41.72***	-42.11***
	(6.183)	(6.954)	(6.273)	(6.254)
Day Care Attendance		4.789***	2.054***	1.605**
		(0.638)	(0.537)	(0.537)
Socioeconomic Status Index			33.87***	32.67***
			(1.254)	(1.231)
Female				2.269
				(1.770)
Constant	514.0***	503.1***	500.7***	509.2***
	(1.457)	(2.712)	(2.377)	(7.798)
Year FEs	No	Yes	Yes	Yes
School Controls	No	No	No	Yes
Observations	11271	9292	9274	9126
R^2	0.030	0.027	0.122	0.129

Note: This table reports the math test score gaps of the students who are categorised as first generation immigrants or second generation immigrants relative to native students. Model (1) reports the unconditional regressions, model (2) add controls for years attended to day care and year fixed effects, model (3), in addition, adds controls for socioeconomic status index and gender, finally, model (4) also includes controls for school characteristics.

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Focusing now on Table 5, attendance to day care is positively and significantly correlated with test performance. In fact, as it can be seen in column 4, the specification that conditions for all the control variables available, an additional year attending day care is associated with an increase in 1.605 PISA points in the math test. However, despite this positive relation between day care attendance and test performance, attending to day care appears to not be important in explaining the test scores gap between native and second-generation students but, it does account for part of the gap with the first-generation immigrants students. An explanation for this results may be related to the fact that, contrary to what expected, second generation students seem to be spending, on average, more years in day care than natives. Thus, as literature also finds, day care seems to be an important feature for test performance regardless of your immigrant status- possibly due to dynamic complementarities- but, due to the fact that there is no different exposure to this factor between the two groups, day care attendance seems to not be explaining the difference in performance between natives and second generation students. Finally, regarding the gap between natives and first generation, a plausible explanation that day care attendance could be capturing is the beneficial effects that an earlier exposition to Finnish institutions may have for students. Nevertheless this result could not be entirely true if the variable "Number of Years Attended in Day Care" also accounted for years attended in the country of origin. As there is no possible access to such information with the data available, in order to test the hypothesis of the beneficial effects of earlier exposure to country specific institutions, the next section analyzes the effects that age of arrival have in explaining the gap.

Table 6: Reading outcomes and day care attendance (2015-2018)

	(1)	(2)	(3)	(4)
	Reading Test Score	Reading Test Score	Reading Test Score	Reading Test Score
Second-Generation	-62.07***	-63.18***	-43.52***	-44.44***
	(8.482)	(7.997)	(7.381)	(7.574)
First-Generation	-109.3***	-88.98***	-73.68***	-71.93***
	(7.157)	(7.872)	(7.806)	(7.767)
Day Care Attendance		5.314***	2.726***	0.873
		(0.690)	(0.623)	(0.605)
Socioeconomic Status Index			32.30***	30.87***
			(1.432)	(1.289)
Female				47.59***
				(1.836)
Constant	529.5***	520.0***	517.7***	503.7***
	(1.646)	(2.931)	(2.673)	(7.999)
Year FEs	No	Yes	Yes	Yes
School Controls	No	No	No	Yes
Observations	11271	9292	9274	9126
R^2	0.043	0.036	0.102	0.177

Note: This table reports the reading test score gaps of the students who are categorised as first generation immigrants or second generation immigrants relative to native students. Model (1) reports the unconditional regressions, model (2) add controls for years attended to day care and year fixed effects, model (3), in addition, adds controls for socioeconomic status index and gender, finally, model (4) also includes controls for school characteristics.

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.3 Age of arrival

In order to continue investigating the effects than an earlier exposition to Finnish institutions may have on improving the test-performance of newly arrived kids, this subsection is going to examine whether the age at which the immigrant kids arrive to Finland is able to explain part of the existing gap between native and immigrants in terms of test score performance. Previous result in Finland, in particular Ansala et al. (2019), find that age at arrival affects educational

attainment in an economically and statistically significant manner. Given that test performance at the age of 15 can be considered a good predictor of future educational attainment, specially for secondary school, finding similar results would be a good sign of my estimates pointing to the right direction. Nevertheless, it is important to take into account that results need to be interpreted carefully. It is still possible than some families may take into account the age of their children before moving to Finland, trying to benefit their sons and daughters in terms of future educational outcomes.

Table 7: Math test scores and age of arrival (2012-2015-2018)

	(1)	(2)	(3)	(4)
	Math Test Score	Math Test Score	Math Test Score	Math Test Score
Second-Generation	-59.65***	-58.94***	-41.86***	-43.48***
	(5.530)	(5.497)	(5.065)	(4.819)
First-Generation	-79.95***	-46.06***	-31.81***	-31.90***
	(4.885)	(9.703)	(9.270)	(9.073)
Age of arrival		-3.924***	-3.108**	-3.118**
		(1.105)	(1.065)	(1.059)
Socioeconomic Status Index			33.55***	32.30***
			(0.985)	(0.917)
Female			1.493	1.103
			(1.318)	(1.326)
Constant	517.4***	523.3***	509.6***	509.2***
	(1.177)	(1.866)	(1.929)	(7.284)
Year FEs	No	Yes	Yes	Yes
School Controls	No	No	No	Yes
Observations	19906	19889	19859	19563
R^2	0.031	0.034	0.129	0.135

Note: This table reports the math test score gaps of the students who are categorised as first generation immigrants or second generation immigrants relative to native students. Model (1) reports the unconditional regressions, model (2) add controls age of arrival and year fixed effects, model (3), in addition, adds controls for socioeconomic status index and gender, finally, model (4) also includes controls for school characteristics.

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7 reports the results for the math test using the baseline specification but with the inclusion of age of arrival as a control. The inclusion of age at arrival in column 2 seems to explain approximately 40% of the gap, however, it is likely that the coefficient is biased upwards. In the following columns, when controls are progressively added for background and school characteristics, it can be observed that the coefficient for age of arrival reduces, although it remains high and statistically significant. Estimates suggest that immigrants who arrived at the age of 0 score on average 31.90 less PISA points meanwhile, those ho arrive when they are 15, score 31.90+15x3.118=78,67 less PISA points on average. On Table 8 the same analysis as in Table 7 is conducted but using the reading scores as the outcome variable. Result suggest a similar pattern as when using the math scores as an outcome variable but with bigger differences between those who arrive as toddlers to those who arrive later. In the case of the reading scores it seams that age of arrival accounts for the 60% of the gap and, while those who arrive at the age of 0 score on average 29.52 PISA points less, arriving at the age of 15 is translated into having an average gap of 29.95+15x7.089=136,28 PISA points.

An important feature of this results is that the coefficient of age of arrival does not change much across specifications. This result may suggest that age of arrival is an important variable in explaining the gap performance for the first generation of immigrants. Moreover, the difference in the coefficient of age at arrival between the two outcome variables suggests an interesting pattern: a later arrival to Finland is more harmful for the subjects that require more country specific skills.

Table 8: Reading test scores and age of arrival (2012-2015-2018)

	(1)	(2)	(3)	(4)
	Reading Test Score	Reading Test Score	Reading Test Score	Reading Test Score
Second-Generation	-63.64***	-63.46***	-46.14***	-50.92***
	(6.386)	(6.367)	(5.763)	(5.405)
First-Generation	-110.0***	-43.26***	-28.08**	-29.52**
	(5.551)	(10.42)	(10.50)	(10.45)
Age of arrival		-7.922***	-7.138***	-7.089***
		(1.195)	(1.202)	(1.198)
Socioeconomic Status Index			32.39***	30.34***
			(1.094)	(0.981)
Female			51.32***	50.72***
			(1.415)	(1.412)
Constant	529.6***	529.7***	492.0***	489.3***
	(1.325)	(2.227)	(2.219)	(8.519)
Year FEs	No	Yes	Yes	Yes
School Controls	No	No	No	Yes
Observations	19906	19889	19859	19563
R^2	0.040	0.042	0.187	0.195

Note: This table reports the reading test score gaps of the students who are categorised as first generation immigrants or second generation immigrants relative to native students. Model (1) reports the unconditional regressions, model (2) add controls age of arrival and year fixed effects, model (3), in addition, adds controls for socioeconomic status index and gender, finally, model (4) also includes controls for school characteristics.

Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

To better understand how the gap evolves for students arriving at different ages, Figures 5 and 6 plot the non-parametric version of the age at arrival gap for both math and reading test scores. The orange line represent the gap between first generation immigrants and natives, while the green line represents the gap between the first generation of immigrants and the second generation of immigrants. First of all it is important to note that no difference is found between being born in Finland and having parents who are immigrants and, arrive before the age of 5 to Finland. Point estimate suggest that the differences between newly arrived kids

and offspings of immigrants who already reside in Finland start to matter if arriving after the age of 5. Interestingly, for both outcomes the negative impact increases for those who arrive between the age of 6 and 8 but, while for the math outcomes no additional impact is found for those arriving up to the age of 11 and a new drop is found for those arriving after the age of 12, for the reading outcomes the gap keeps increasing the later the arrival.

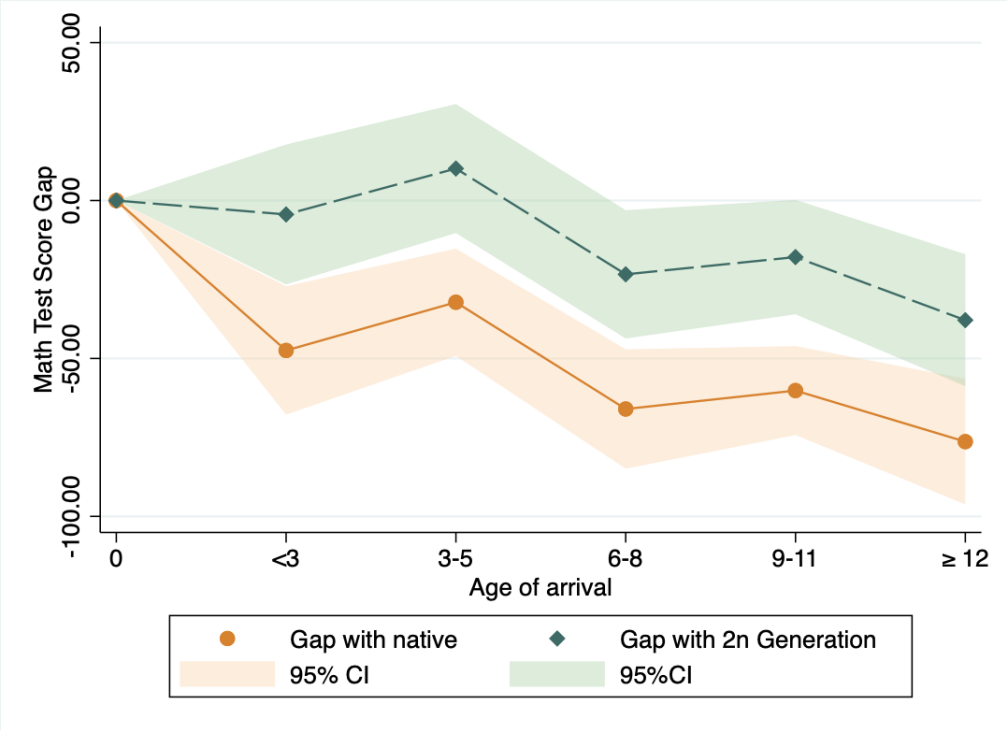

Figure 5: Fist generation age of arrival math scores gap estimates against native and second generation students.
Note: Points estimates and 95% confidence interval are plotted for the specification that includes both background characteristics and school controls

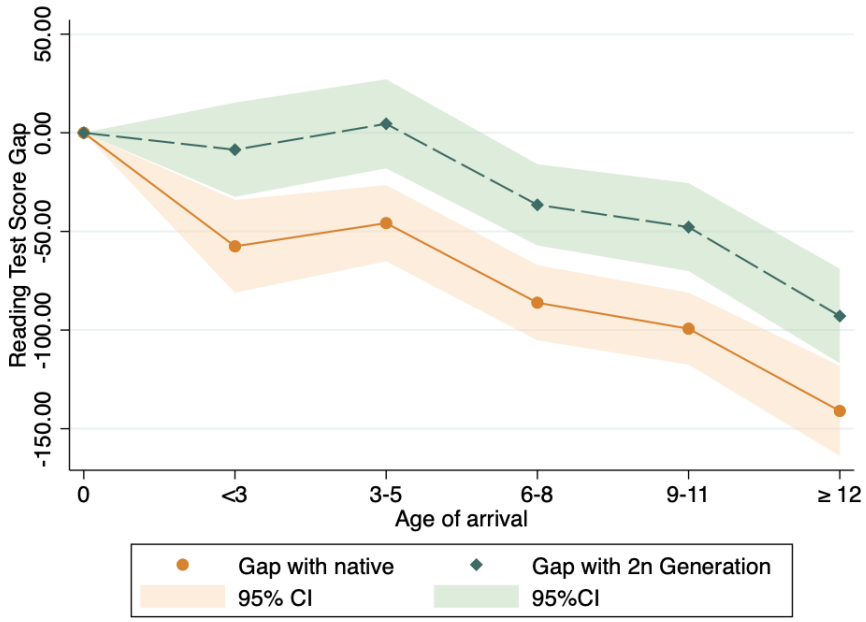

Figure 6: Fist generation age of arrival reading scores gap estimates against native and second generation students.
Note: Points estimates and 95% confidence interval are plotted for the specification that includes both background characteristics and school controls

Turing the attention to the gap between the first generation of immigrants and natives, similar to Ansala et al. (2019), a large negative impact is found for those arriving before the age of three compared to being born in Finland, but no additional impact of arriving between 3-5. Also, estimates suggest that age of arrival starts to matter again when arriving being older than 5. Similar to the second generation first generation gap, it is found that for the math test scores there is no additional impact for those arriving until the age of 11 and a small drop for those arriving after 12. In the case of the reading literacy test score gap, the gap continuously increases with age of arrival.

Putting together this results with previous ones, they suggest that age of arrival has indeed a causal negative impact on educational outcomes. In my case, there seems to be a negative relation between standardized tests performance and age of arrival. In particular the estimates suggest two critical ages: arriving before the age of 3 and arriving at the age between 6-8.

Understanding the mechanisms through which age of arrival affects educational outcomes is an essential need in order to design proper integration policies and, although I don't have a strong research design to study the channels through which age of arrival impacts test performance, my results, and the pattern they display, seem to be more consistent with some explanation that others.

Among the different mechanisms that literature has proposed, such as network disruption, parental exposure to the country, children own exposure to Finnish institutions and the presence of country specific skills such as language, I believe that the combination of the last two is the one that suits my results the best. Taking into account the critical ages at which an increase in the gap is observed, before 3 and after 5, it does coincide with the ages at which parents can decide weather they want to put their kids in day-care. The reason why arriving before 3 in Finland does not make a difference in terms of educational outcomes may be due to the fact that exposition to Finnish institutions does not change much between those who arrive before the age of 3 or 5. On average first generation of immigrants spent 2.93 years in day care on my sample, suggesting they start between the ages of 3 and 4. In this sense, for example, country specific skills such as language, contrary to natives, can not be learned in their households and hence, the country were they are until they do enter in day care does not has much effect on future educational outcomes in Finland. For those who arrive after they have turned 6, the drop in their test performance may be related to the fact that they miss the possibility to attend to Finnish day care. This has a negative impact of the general performance of first generation immigrants compared to their peers. A possible explanation of this negative impact can be found in the fact that the quality of the institutions in the source country are worse than the Finnish ones or, because they miss the chance to acquire a whole set of country specific skill that all other students acquire before entering to compulsory school. However, although the

hypothesis about the exposure to early institutions such as day care and the existence of country specific skills seem to fit the pattern displayed in both figures 5 and 6, no strong conclusion can be made about the mechanisms through which age at arrival affects future school outcomes.

4.4 Heterogeneity by Socioeconomic status

One of the main issues when studying the native-immigrant gap is that both natives and immigrant population are not necessarily homogeneous. For this reason, this section analyzes how the native-immigrant test score gap evolves across my sample when divided in percentiles by socioeconomic status. Figures 7 and 8 plot the point estimates of equation 1 when being computed for each percentiles using the math test scores as outcome variable[1]. The particularity in this section is that it includes the language spoken at home as a control variable. Unfortunately, PISA does not contain precise information about the source country of the students or the parental country of birth and hence, the language they speak at home is the best proxy available. My intention is to capture the fact that immigrants in different percentiles come from different regions. For example, it it likely that those immigrants who are on the top of the socioeconomic status distribution come from other European countries while those at the bottom come from African and Middle-East countries.

The unconditional estimates for the first generation gap reveals no clear heterogeneity across percentiles (Table 7). Only for those students in forty percentile and on the ninety percentile the gap seems to be smaller, around 30 PISA points. For the remaining percentiles, the unconditional gap is situated in between 60 and 75 PISA points. Regarding the conditional gap, it becomes substantially smaller just in the top percentiles. In fact, it becomes statistically insignificant for those it the top three percentiles. On the top parts of the socioeconomic distri-

[1]Figures 10 and 11 in the Appendix estimate the same specification but using the reading test scores as outcome variable

bution, the gap seems to be explained in part due to the differences in the different languages spoken at home (Table 9 in the Appenix). However, claims can't be directly made about the causal effect of language or if its capturing the effect of a particular source country first, because my estimates are very imprecise and secondly, because language spoken at home may be reflecting other unobservable factors that may affect test performance. Regarding the students on the rest of the distribution, there seems to persist a negative gap in terms of test performance also after the inclusion of the controls.

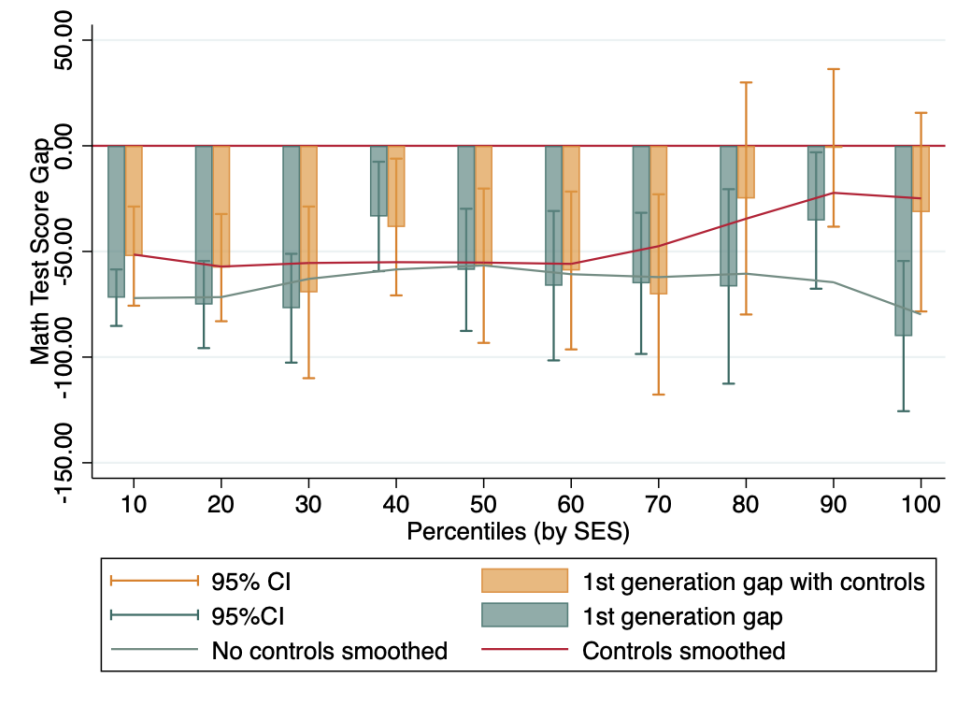

Figure 7: First generation immigrant-native gap by percentiles in math test scores (native as baseline category).
Note: This figure plots the point estimate and the 95% confidence interval of the main specification by percentiles together with a line that smooths the gap distribution across percentiles.

In Table 8, the unconditional estimates also reveal no clear pattern in the test score gap for the second generation. However, for every percentile, except from the 90th, second generation immigrants significantly trail natives in terms of test performance. The conditional gaps are substantially smaller and seem to display a slight U-shaped pattern. In fact, for those in the top three percentiles the 60th and the 20th, it can't be rejected the null hypothesis that the

gap is statistically significant different from 0 and, for those students in the 10th and 30th percentile the gap is substantially smaller than those on the 40th, 50th or 70th. This results suggest, specially for those students at the tails of the distribution, that test performance can be largely explain by language spoken at home or school characteristics. However, similarly to the results found for the first generation, claims can't be made about the exact feature that language spoken at home is capturing.

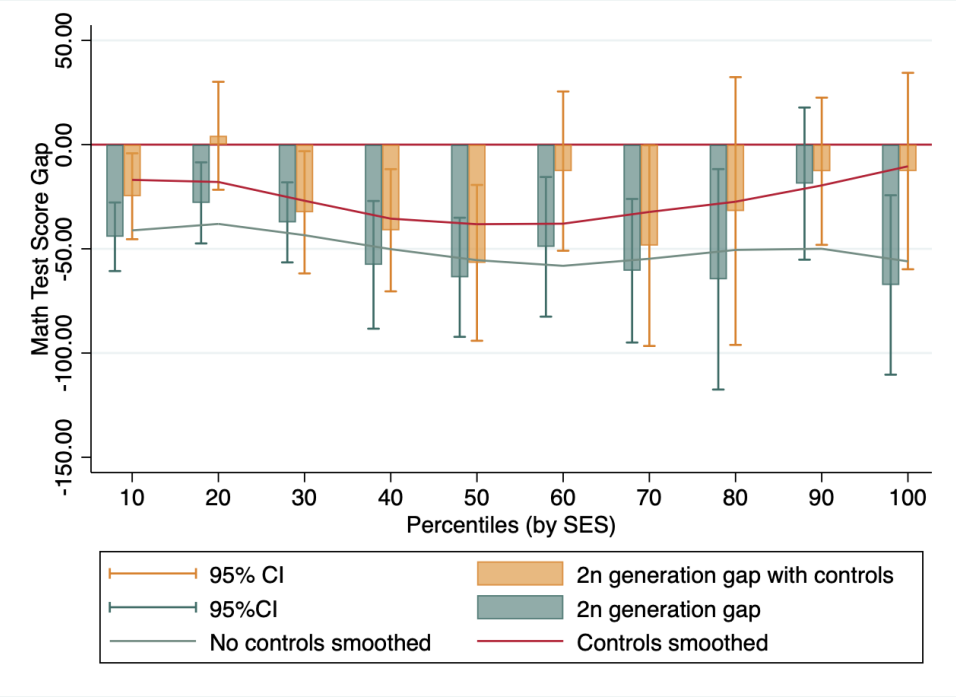

Figure 8: Second generation immigrant-native gap by percentiles in math test scores (native as baseline category).
Note: This figure plots the point estimate and the 95% confidence interval of the main specification by percentiles together with a line that smooths the gap distribution across percentiles.

Overall in this section, results suggest that the gap in the top parts of the distribution disappears on the conditional specification regardless of you immigrant status. However, it is unclear to which factors the gap can be attributed to as it is unclear what the language spoken at home is indeed capturing. Also, it is worth realizing that even in the conditional specifications large differences remain between native and immigrants on the middle part of the distribution and on the bottom part for the fist generation of immigrants. Even though my estimates are also imprecise in this percentiles, it could be interesting to further inquire in the gap across a

heterogeneous population such as the immigrant one. This will allow to better target specific groups of the population and design more efficient policies so that no student is left behind.

4.5 Heterogeneity by Language Spoken at Home

Next, this section examines how differences in test performance differs across groups of students who speak a different language than Finnish at home. As already mentioned, point estimates in this section do not report the immigrant native gap because first, not all natives speak Finnish at their homes and second, because some students with immigrant background speak Finnish at their homes. Hence, in this section the parameters of interest report the gap between those who speak a specific language compared to those who speak Finnish. The unconditional gap is reported in orange while the specification that adds the controls for both background and school characteristics is reported in green. Figure 9 reports the language gap for the math scores and Figure 10 report the language gap for the reading scores.

As expected, the test scores gap is larger in the reading specification than in math one. However, both of them display a similar distribution and a substantial heterogeneity. First of all, the gap is much smaller for those students that speak a language of one the language neighboring countries to Finland such as Estonian, Swedish or Russian. Nevertheless, both students who speak Estonian and Russian at home perform around 0.3 standard deviations lower than those who speak Finnish in the math test. In the unconditional specifications, those who speak Somali are the ones who perform the worst, relative to the Finnish speakers, in the the math tests while in the reading test, those who speak Arabic are the ones that have the lowest relative score. However, it needs to be pointed out that the gaps are extremely large- more than one standard deviation for the math test scores- for those students whose main language is Somali, Turkish, Kurdish, Thai of Chinese. The conditional gaps still display a similar patter, although the

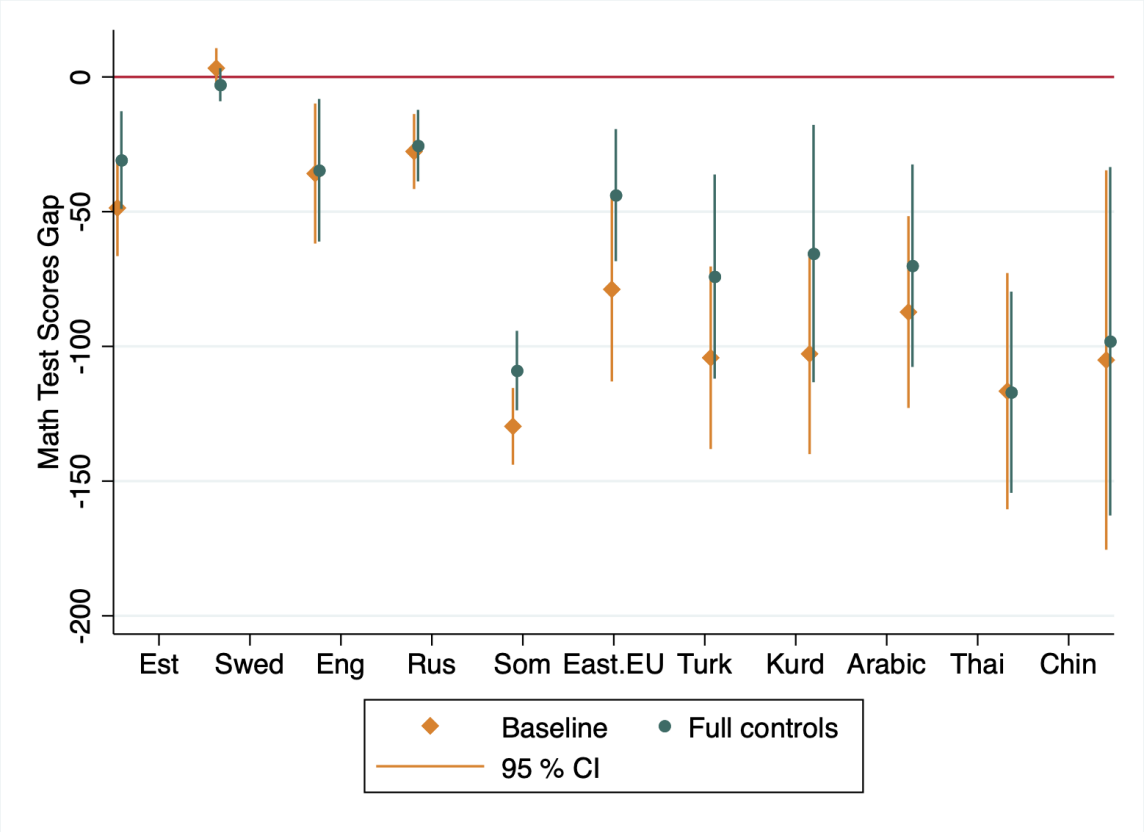

Figure 9: Math test scores gap by language spoken at home (Finnish as baseline category).
Note: The x-axis display the language to which Finnish is compared. Est stands for Estonian, Swd for Swedish, Rus for Russian, Som for Somali, EstEU groups Romany and Albanian, Turk for Turkish, Kurd for Kurdish and Chin for Chinese

gaps become substantially smaller for those who speak either Romany or or Albanian (Est.Eu), Turkish, Kurdish or Arabic and remain virtually unchanged for those who speak Swedish, English, Russian, Thai or Chinese.

These results seem to confirm the intuition that there exist a huge variation in terms of test performance among the population. However, it is important to read them with caution, as they are quite imprecise as the sample available is unlikely to be representative of the full population speaking a certain language. Moreover, it is also important to remember that point estimates could be capturing other features rather than country of origin. For example, those who speak Somali or Kurdish, what point estimates may be suggesting is that those kids who arrive as refugees, or are the offspring of the refugees, face a much harder integration process and this one is reflected in their school performance.

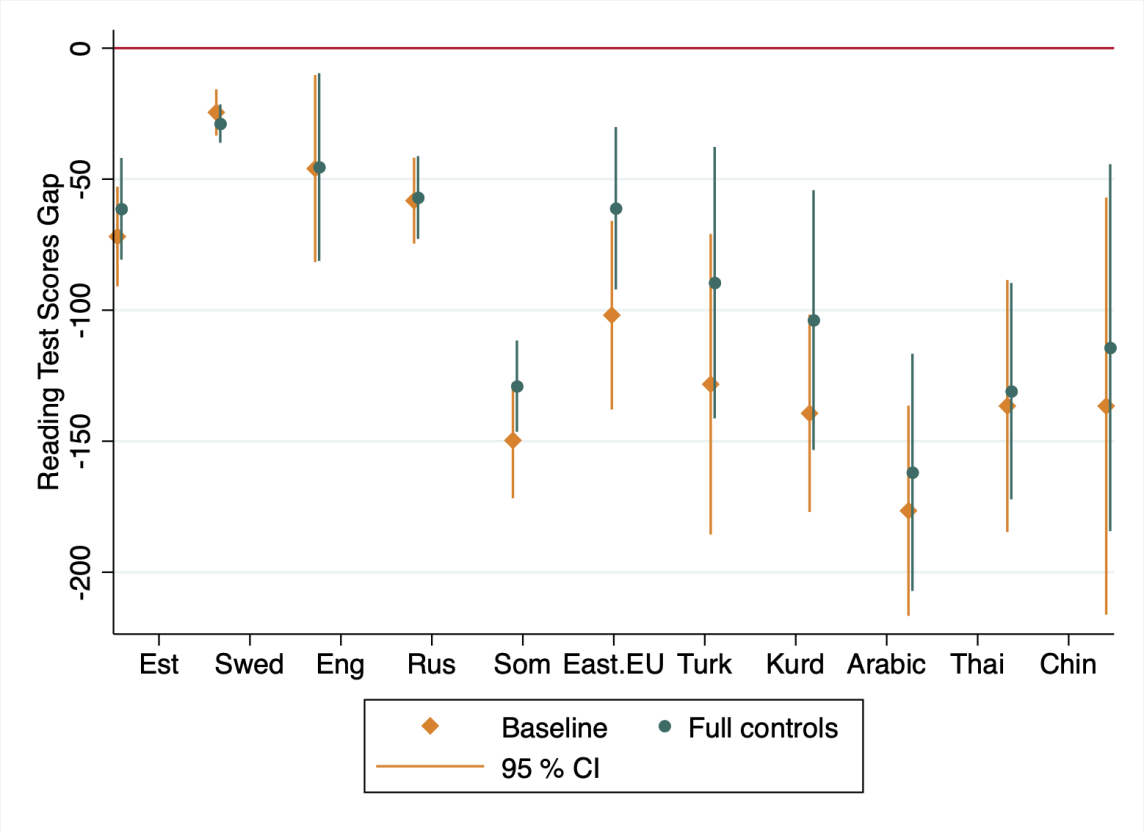

Figure 10: Reading test score gap by language spoken at home (Finnish as baseline category). *Note:* The x-axis display the language to which Finnish is compared. Est stands for Estonian, Swd for Swedish, Rus for Russian, Som for Somali, EstEU groups Romany and Albanian, Turk for Turkish, Kurd for Kurdish and Chin for Chinese

5 Conclusions

This paper has studied how the immigrant students at the age of 15 have performed in the recent past compared to its native peers in Finland. It has been found that unconditionally, first generation of immigrants perform around 0.75 standard deviation lower in the math tests and more than 1 standard deviation lower in the reading literacy tests while, second generation immigrants perform around 0.5 standard deviations lower in the math tests and 0.6 standard deviations lower in the reading literacy tests. Also it has found that this gap is fairly reduced, although not completely offset once taken into account certain variables.

First, estimates indicate that socioeconomic status plays an important role in explaining part of the test scores gap, suggesting that school performance is correlated with childhood resources.

Second, a limited effect of day care attendance to the size of the gap has been found. This result, contrary to what expected, is more likely to be explained due to measurement error or the representatives of the sample. More detailed research should be conducted in disentangling the effects of day care attendance on future outcomes as it has also been found that the reading score gap, which account for more country specific skills, is always bigger than the math score gap. Third, in line with previous literature we have found an important effect of age of arrival in explaining the gap for the first generation. Critical ages for arrival appear to be before the age of three and between 6-8 years old. In this sense, it would be interesting to further investigate the mechanisms through which age of arrival affects test performance. Results also suggest that the test score gap is heterogeneous across the the sample of immigrants. The conditional gap is much smaller for all those situated in the top percentile of the distribution and those who speak a language that can be placed in Europe.

Taken into account this results, general policies targeting disadvantage population may be helpful in reducing the gap. On average, immigrant population have a lower socioeconomic status and part of the gap is explained by this lower access to resources rather than their immigrant status. However, my results also suggest that some specific policies, targeting certain profiles of immigrants such as those arriving at later stages of the basic education, may be necessary to efficiently reduce the gap and make the integration process much smoother. In this sense, future research should focus on understanding the causal role that age of arrival has on future outcomes and how it interacts with other aspects such as exposure to Finnish institutions or parental integration process, but not just studying the native-immigrant gap from an individual perspective, but also taking into account structural factors such as intuitions.

6 References

Ansala, L., Hämäläinen, U., Sarvimäki, M. (2019). "Age at arrival, parents and neighborhoods: understanding the educational attainment of immigrants' children". *Journal of Economic Geography*, forthcoming.

Ansala, L., Aslund, O., Sarvimäki, M. (2018). "Immigration hisotry, entry jobs, and the labour market integration of immigrants". *Working Paper*

Aslund, O., Böhlmark, A., Skans, O.N. (2015). "Childhood family experiences and the social integration of young migrants". *Journal of Labour Economics*, 35:1416-1440.

Berlinski, S., Galiani, S., Manacorda, M. (2008). "Giving children a better start: Preschool attendance and school-age profiles". *Journal of Public Economics*, 92:1416-1440.

Böhlmark, A. (2008). "Age at immigration and school performance: A siblings analysis using Swedish register data". *Labour Economics*, 15(6):1366-1387.

Borjas, G. (1995). "Ethnicity, Neighborhoods, and Human-Capital Externalities". *The American Economic Review*, 85(3):365-390.

Borgna, C. (2016):" Multiple paths to inequality. How institutional contexts shape the educational opportunities of second-generation immigrants in Europe", *European Societies*, 18(2): 180-199.

Cameron, S.V., Heckman, J.J. (2001):"The dynamics of educational attainment for black, hispanic and white males",*Journal of Political Economy*, 109: 455-499.

Card, D.(2005):"Is the new immigration really so bad?",*Economic Journal*, 115: 300-323.

Card, D., Schmidt, C. (2003):"Editorial: Symposium on Second-Generation Immigrants and the Transition to Ethnic Minorities",*Journal of Population Economcs*, 16(4): 707-710.

Chmielewski (2019). "The Global Increase in the Socioeconomic Achievement Gap, 1964 to 2015". *American Sociological Review*, 84(3):517-544.

Cornelissen, T., Dustmann,C. (2019). "Early School Exposure, Test Scores, and Noncognitive Outcomes." *American Economic Journal: Economic Policy*,11(2):35-63.

Cornelissen, T., Dustmann C., Raute A., and Schönberg U. (2018). "Who Benefits from Universal Child Care? Estimating Marginal Returns to Early Child Care Attendance." *Journal of Political Economy*, 126(6):2356-2409.

Cuhna, F., Heckman, J. (207). "The Technology of Skill Formation." *American Economic Review*, 97(2):31-47.

Currie, J., Almond, D. (2011). "Human capital development before age five." *Handbook of Labor Economics*,vol. 4B:1315-1486.

Dearden, L., Machin, S., Reed, H. (1997). " Intergenerational Mobility in Britain." *The Economic Journal*, 107(440):47-66.

Drange, N., Havnes, T. (2019)."Early Childcare and Cognitive Development: Evidence from an Assignment Lottery." *Journal of Labor Economics*, 37(2): 581-620.

Dustman, C., Frattini, T., Lanzara, G. (2012)."Educational achievement of second-generation immigrants: an international comparison." *Economic Policy*, 27(69): 143-185.

Haveman, R., Wolfe, B. (1995)."The Determinants of Children's Attainments: A Review of Methods and Findings." *Journal of Economic Literature*, 33(4): 1829-1878.

Jänti, M., Røed,K., Naylor, R., Björklund, A., Bratsber, B., Raaum, O., Österbacka, E., Eriksson, T. (2006)."American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic countries, the United Kingdom and the

United States." *IZA Institute of Labor Economics*, Discussion Paper No. 1938.

Ruhm, C., Waldfogel, J. (2012)." Long-term effects of early childhood care and education." *Nordic Economic Policy Review*, 1(1): 23-51.

Organisation for Economic Co-operation and Development. (2009). "PISA Data Analysis Manual: SPSS, Second Edition." Paris, France: *OECD Publishing*.

Organisation for Economic Co-operation and Development. (2010). "Strong performers and successful reformers in education: Lessons from PISA for the United States". Paris, France: *OECD Publishing*.

Ours, J., Veenam, J. (2003)."The educational attainment of second-generation immigrants in The Netherland." *Journal of Population Economics*, 16: 739-753.

Pekkarinen, T., Uusitalo, R., Kerr, S. (2009). "School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform", *Journal of Public Economics*, 93(7-8): 965-973.

Salmela-Aro, K., Chmielewski, A. (2019). "Socioeconomic Inequality and Student Outcomes in Finnish Schools". 10.1007/978-981-13-9863-6_9.

Schneeweis, N. (2011). "Educational institutions and the integration of migrants", *Journal of Population Economics*, 24(4):1281-1308

Schepf, S.V. (2007). "Immigrants' educational disadvantage: an examination across ten countries and three surveys", *Journal of Population Economics*, 20(3):527-545

7 Appendix

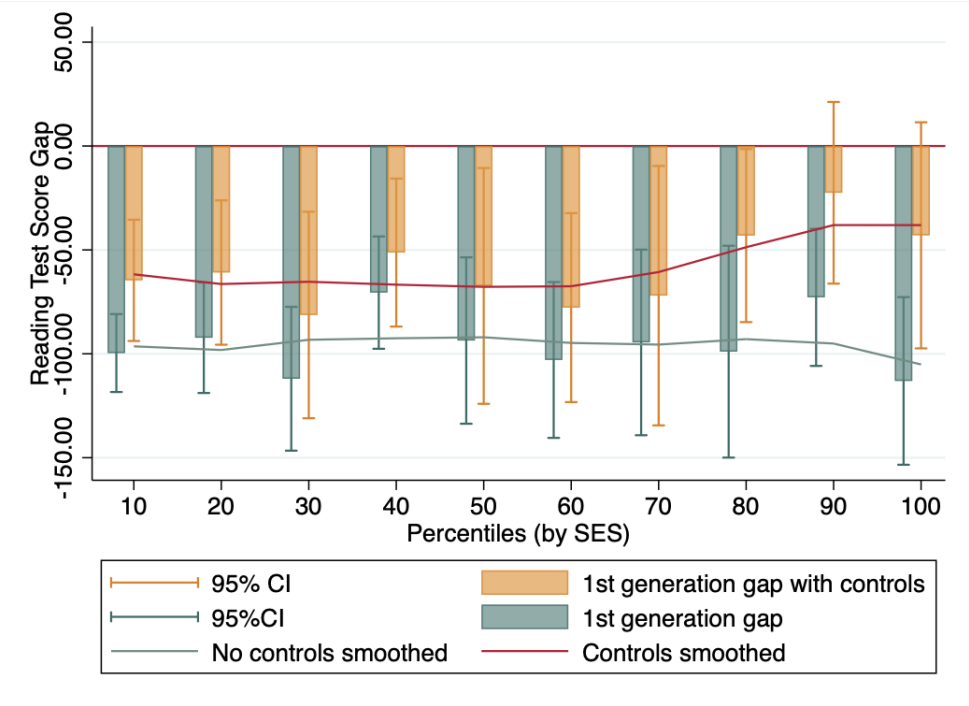

Figure 11: First generation immigrant-native gap by percentiles in reading test scores (native as baseline category).
Note: This figure plots the point estimate and the 95% confidence interval of the main specification by percentiles together with a line that smooths the gap distribution across percentiles.

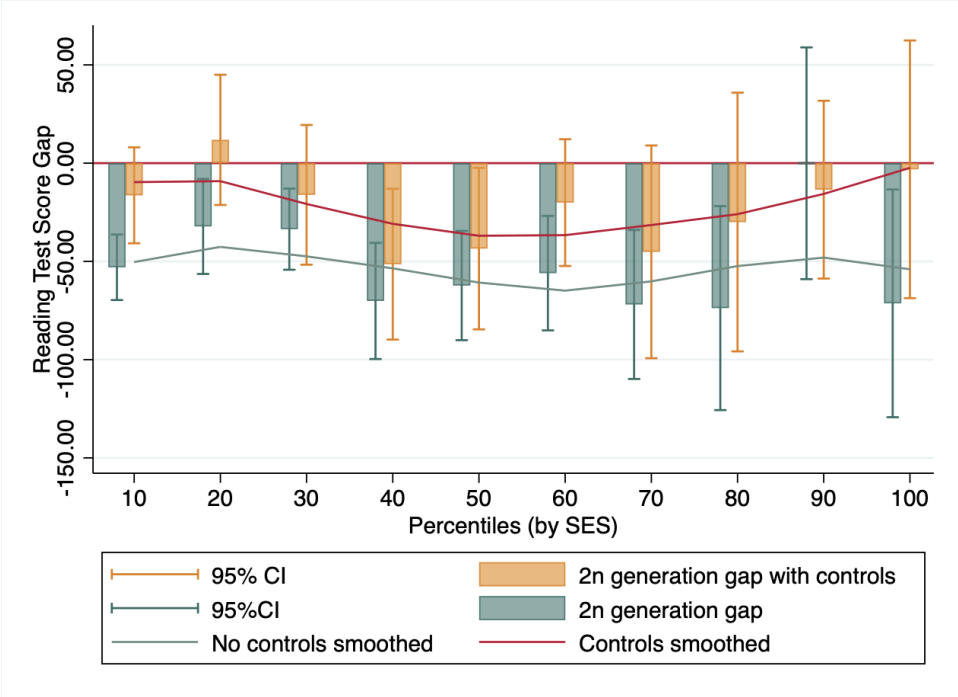

Figure 12: Second generation immigrant-native gap by percentiles in reading test scores (native as baseline category).
Note: This figure plots the point estimate and the 95% confidence interval of the main specification by percentiles together with a line that smooths the gap distribution across percentiles.

Table 9: Percentiles no controls

	(1) Math Test Score	(2) Math Test Score	(3) Math Test Score	(4) Math Test Score	(5) Math Test Score	(6) Math Test Score	(7) Math Test Score	(8) Math Test Score	(9) Math Test Score	(10) Math Test Score
Second-Generation	-44.22***	-27.96**	-37.31***	-57.68***	-63.65***	-49.04**	-60.51***	-64.63*	-18.71	-67.34**
	(8.374)	(9.898)	(9.776)	(15.59)	(14.55)	(17.05)	(17.53)	(26.91)	(18.58)	(21.90)
First-Generation	-71.89***	-75.11***	-76.87***	-33.45*	-58.70***	-66.21***	-65.11***	-66.56**	-35.37*	-90.05***
	(6.799)	(10.50)	(13.10)	(13.18)	(14.71)	(17.99)	(17.00)	(23.42)	(16.42)	(18.09)
Constant	473.3***	489.7***	495.9***	502.6***	511.7***	517.9***	529.4***	538.9***	555.1***	555.5***
	(2.621)	(2.005)	(2.218)	(2.137)	(2.049)	(2.130)	(2.028)	(2.164)	(2.120)	(2.209)
Observations	2000	2008	1991	2026	1996	2001	2001	2035	1971	1877
R^2	0.061	0.030	0.033	0.014	0.017	0.018	0.017	0.015	0.004	0.025

Note: This table reports the math test score gaps by socioeconomic status percentiles of the students who are categorised as first generation immigrants or second generation immigrants relative to native students. All columns report the unconditional gap and the regressions are run separately for each percentile. In ascendant order, each column stands for one percentile.
Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 10: Percentile and controls

	(1) Math Test Score	(2) Math Test Score	(3) Math Test Score	(4) Math Test Score	(5) Math Test Score	(6) Math Test Score	(7) Math Test Score	(8) Math Test Score	(9) Math Test Score	(10) Math Test Score
Second-Generation	-24.77*	4.237	-32.48*	-41.09**	-56.73**	-12.72	-48.46*	-31.87	-12.78	-12.69
	(10.49)	(13.18)	(14.92)	(14.92)	(19.03)	(19.45)	(24.50)	(32.69)	(17.97)	(23.97)
First-Generation	-52.21***	-57.67***	-69.37***	-38.45*	-56.76**	-59.02**	-70.35**	-24.94	-0.996	-31.40
	(11.94)	(12.91)	(20.67)	(16.47)	(18.59)	(19.01)	(24.12)	(27.94)	(18.98)	(23.91)
Female	0.263	2.229	6.118	1.649	2.836	1.483	0.606	-2.755	-10.67**	1.295
	(3.919)	(3.660)	(3.868)	(3.787)	(3.912)	(3.964)	(3.949)	(3.828)	(3.611)	(3.980)
Estonian	3.805	-42.06	77.82***	144.4			137.5***	141.6***		-89.94
	(33.79)	(39.61)	(22.72)	(80.23)			(29.78)	(28.17)		(66.10)
Somali	-78.41*	-131.2***	22.19	30.37	-109.9**	-121.4***	1.513	51.96	-124.8*	-69.04
	(31.44)	(33.26)	(19.01)	(79.24)	(39.51)	(28.87)	(12.35)	(33.68)	(51.30)	(42.81)
Albanian	31.02	-49.53	54.27**		-125.5***	-13.46	95.87***			-11.66
	(28.13)	(31.36)	(20.07)		(31.39)	(27.16)	(4.360)			(64.55)
English	-74.05	-55.54	94.39**	82.28	-82.35*	-115.8*	72.26	231.2***	-34.10	-8.991
	(48.59)	(44.48)	(34.84)	(79.23)	(37.35)	(54.59)	(41.83)	(30.59)	(37.40)	(31.84)
Chinese	-84.51	-184.2***	56.74**	131.8	-285.5***		-111.2	179.2**	39.59	-146.7**
	(47.75)	(47.57)	(20.14)	(76.87)	(61.23)		(72.53)	(54.52)	(53.05)	(50.97)
Turkish	-13.93	-97.01	64.99*	178.7	-127.2***	-183.6***	39.46**	3.230	-202.2***	-110.6**
	(35.98)	(59.77)	(26.44)	(91.09)	(28.85)	(27.99)	(12.55)	(31.98)	(36.13)	(37.47)
Romany	-83.65*	-170.7***	74.09**	133.2	-143.8***			244.8***	45.60	21.59
	(41.97)	(48.74)	(23.29)	(79.66)	(29.46)			(31.17)	(52.16)	(34.01)
Finnish	-5.935	-56.27	103.0***	109.1	-85.49**	-25.82	72.06**	156.5***	-16.85	68.81*
	(28.56)	(34.57)	(19.05)	(79.36)	(29.01)	(27.90)	(26.43)	(30.53)	(36.29)	(34.41)
Swedish	-3.232	-64.70	102.1***	110.4	-88.64**	-15.67	61.20*	148.9***	-26.46	66.91
	(30.87)	(36.00)	(20.63)	(79.26)	(30.24)	(28.86)	(26.91)	(31.17)	(35.36)	(35.21)
Russian	13.45	-40.07	115.0***	163.8*	-87.73**	-7.871	85.74*	156.9***	-76.63	33.67
	(34.69)	(33.56)	(20.24)	(78.61)	(27.64)	(35.36)	(35.22)	(34.55)	(43.25)	(30.59)
Arabic	-37.03	-128.0***	154.2***		-95.80**	-0.452	-4.957	35.03		-17.57
	(31.84)	(31.89)	(26.87)		(31.54)	(20.42)	(27.42)	(39.22)		(33.95)
Sami	-111.1**	-215.3***	-9.604				139.6***			104.3*
	(35.90)	(33.23)	(19.59)				(30.54)			(47.22)
Thai	-92.44**	-70.84	19.60	-60.00	-215.1***	-111.5***	-134.6***	111.4***		-32.50
	(32.52)	(37.89)	(38.68)	(78.42)	(29.07)	(20.70)	(38.92)	(31.15)		(44.91)
Another language (FIN)	-32.80	-97.60**	123.4***	70.56	-72.20*	-72.26**	57.14*	94.81***	-66.25	-0.548
	(29.26)	(32.13)	(20.26)	(78.08)	(29.25)	(23.12)	(23.91)	(26.11)	(39.79)	(36.62)
Invalid	-18.50	-141.0**	65.72***	23.27	-122.4*	-117.6***	-22.26	8.478	-17.49	-21.10
	(38.15)	(47.49)	(19.74)	(79.39)	(54.20)	(33.65)	(50.21)	(39.96)	(41.13)	(43.93)
No Response	-126.7***	-78.83*	81.47***	126.6	-129.3***	-3.789	20.23	147.2***	-124.2**	-9.558
	(36.66)	(34.86)	(19.57)	(90.88)	(25.47)	(34.89)	(35.05)	(37.99)	(41.32)	(43.41)
Constant	486.1***	510.5***	412.9***	382.0***	583.6***	541.8***	475.1***	383.7***	615.3***	494.8***
	(36.57)	(36.11)	(26.71)	(79.83)	(34.06)	(31.60)	(34.72)	(31.58)	(38.94)	(37.04)
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1970	1989	1967	1999	1964	1969	1974	2000	1929	1840
R^2	0.118	0.062	0.067	0.049	0.039	0.047	0.045	0.047	0.047	0.086

Note: This table reports the math test score gaps by socioeconomic status percentiles of the students who are categorised as first generation immigrants or second generation immigrants relative to native students. All columns report the gap from the regression model that includes both background characteristics and school controls. Regressions are run separately for each percentile. In ascendant order, each column stands for one percentile.
Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.